

**Белорусский государственный университет**  
**Механико-математический факультет**  
**Кафедра дифференциальных уравнений и системного анализа**

**Аннотация к дипломной работе**  
**«Сервис семантической навигации по новостным данным»**

**Мытько Андрей Леонидович,**

**руководитель Перез Чернов Александр Хуанович**

**2015**

Дипломная работа содержит 45 страниц, 5 рисунков, 7 таблиц, 3 приложения, 22 источника.

Ключевые слова: СЕМАНТИЧЕСКАЯ НАВИГАЦИЯ, КЛАССИФИКАЦИЯ, ДЕТАЛИЗАЦИЯ ТЕКСТА, ОБРАБОТКА ЕСТЕСТВЕННОГО ЯЗЫКА, ВЕБ-СЕРВИС, МАСШТАБИРУЕМОСТЬ, SCALA.

Дипломная работа представляет собой описание реализации интерактивного веб-сервиса, осуществляющего возможность классификации документов и детализации текста для обеспечения семантической навигации.

Целью дипломной работы является реализация сервиса семантической навигации по новостным данным на основе масштабируемых технологий.

Для достижения поставленной цели использовались

- техники машинного обучения,
- наивный байесовский классификатор,
- векторная модель признаков TF-IDF.

В дипломной работе представлены следующие результаты:

- 1) рассмотрены существующие подходы обработки естественного языка и примеры похожих сервисов,
- 2) спроектирована архитектура сервиса,
- 3) реализована серверная функциональность,
- 4) построен веб-сервис, осуществляющей семантическую навигацию.

Дипломная работа выполнена автором самостоятельно.

The thesis is presented in the form of an explanatory note of 45 pages, 5 figures, 7 tables, 3 appendices, 22 references.

Keywords: SEMANTIC NAVIGATION, CLASSIFICATION, TEXT DETAILING, NATURAL LANGUAGE PROCESSING, WEB-SERVICE, SCALABILITY, SCALA.

The thesis project examines creation of an interactive web-service, which implements documents classification and text detailing for providing semantic navigation.

The goal of the thesis project is realization of the semantic navigation service over news data based on scalable technologies.

To achieve the goal we used

- machine learning techniques,
- naïve Bayes classification,
- vector space model of features TF-IDF.

The main results of the thesis project are as follows:

- 1) existing natural language processing approaches and some other services alike were considered,
- 2) service architecture was projected,
- 3) server functionality was implemented,
- 4) web-service providing semantic navigation was built.

The thesis was written solely by the author.